

**REMARKS**

In the Office Action, dated September 19, 2007, the Examiner rejects claims 1-9 and 15-19 under 35 U.S.C. § 112, 2<sup>nd</sup> paragraph, as allegedly failing to particularly point out and claim the subject matter that Applicant regards as his invention; and rejects claims 1-29 under 35 U.S.C. § 103(a) as allegedly being unpatentable over U.S. Patent Application Publication No. 2004/0177015 to Galai (hereinafter "GALAI") in view of U.S. Patent No. 6,952,730 to Najork et al. (hereinafter "NAJORK"). Applicant respectfully traverses these rejections.<sup>1</sup>

By way of the present amendment, Applicant cancels claims 2 and 6 without prejudice or disclaimer, amends claim 1 to include the features of claim 6 and to improve form, and amends claims 4, 7-9, 10, 14-16, 18-20, 24, and 25 to improve form. No new matter is being added by way of the present amendment. Claims 1, 3-5 and 7-29 are pending.

Pending claims 1, 3-5, 7-9, and 15-19 stand rejected under 35 U.S.C. § 112, 2<sup>nd</sup> paragraph, as allegedly failing to particularly point out and claim the subject matter that Applicant regards as his invention. Applicant respectfully traverses this rejection.

The Examiner alleges that claims 1-9 recite the feature "the extracted set of URLs" for which there is insufficient antecedent basis (Office Action, p. 2). Without necessarily agreeing with the Examiner, Applicant has amended claims 1, 7, and 9, to address the Examiner's concerns, in order to expedite prosecution.

---

<sup>1</sup> As Applicant's remarks with respect to the Examiner's rejections are sufficient to overcome these rejections, Applicant's silence as to assertions by the Examiner in the Office Action or certain requirements that may be applicable to such rejections (e.g., whether a reference constitutes prior art, motivation to combine references, assertions as to dependent claims, etc.) is not a concession by Applicant that such assertions are accurate or such requirements have been met, and Applicant reserves the right to analyze and dispute such assertions/requirements in the future.

The Examiner alleges that claims 15-19 recite the feature “the extracted URLs” for which there is insufficient antecedent basis (Office Action, p. 2). Without necessarily agreeing with the Examiner, Applicant has amended claims 15, 18, and 19, to address the Examiner’s concerns in order to expedite prosecution.

For at least the foregoing reasons, Applicant submits that claims 1, 3-5, 7-9, and 15-19 meet the requirements of 35 U.S.C. § 112, 2<sup>nd</sup> paragraph. Accordingly, Applicant respectfully requests that the rejection of claims 1-5, 7-9, and 15-19 under 35 U.S.C. § 112, 2<sup>nd</sup> paragraph be reconsidered and withdrawn.

Pending claims 1, 3-5, and 7-29 stand rejected under 35 U.S.C. § 103(a) as allegedly being unpatentable over GALAI in view of NAJORK. Applicant respectfully traverses this rejection.

At the outset, Applicant submits that the Examiner appears to allege that GALAI anticipates claims 1, 2, 3, 5, 11, 12, and 13. The Examiner only relies on NAJORK for disclosing “wherein the comparison of the second URL to the clean set of URL is based on a comparison of a fingerprint value calculated for each of the URLs in the clean set of URLs” (Office Action, p. 4). This feature is not recited in claims 1, 3, 5, 11, 12, and 13. Applicant requests that the Examiner clarify the grounds of rejection for these claims.

Amended independent claim 1 is directed to a method that includes extracting a set of uniform resource locators (URLs) from at least one document, analyzing the set of URLs extracted from the at least one document to determine those in the set of URLs that contain session identifiers by locating the session identifiers in the set of URLs extracted as sub-strings that occur in multiple URLs of a web site, generating a clean set of URLs from the set of URLs extracted from the at least one document by removing the session

identifiers, and determining when at least one particular URL has already been crawled based, at least in part, on a comparison of the particular URL to the clean set of URLs. GALAI and NAJORK, whether taken alone or in any reasonable combination, do not disclose or suggest this combination of features.

For example, GALAI and NAJORK do not disclose or suggest analyzing the set of URLs extracted from the at least one document to determine those in the set of URLs that contain session identifiers by locating the session identifiers in the set of URLs extracted as sub-strings that occur in multiple URLs of a web site, as recited in claim 1.

This feature is similar to the feature previously recited in claim 6. The Examiner did not address this feature (Office Action, p. 3). Thus, a *prima facie* case of obviousness has not been established with respect to amended claim 1. Nevertheless, Applicant submits that neither GALAI nor NAJORK discloses or suggests analyzing the set of URLs extracted from the at least one document to determine those in the set of URLs that contain session identifiers by locating the session identifiers in the set of URLs extracted as sub-strings that occur in multiple URLs of a web site, as recited in claim 1. The Examiner appears to rely on paragraphs 0012, 0013, and 0015 of GALAI for allegedly disclosing “analyzing the extracted set of URLs to determine those in the set of URLs that contain session identifiers” (Office Action, p. 3). Applicant submits that none of these sections, or any other section, of GALAI discloses or suggests the above feature of amended claim 1, where URLs that contain session identifiers are determined by locating the session identifiers in the set of URLs extracted as sub-strings that occur in multiple URLs of a website.

With regard to NAJORK, the Examiner relies on col. 9, lines 4-17 of NAJORK

for allegedly disclosing that the comparison of the second URL to the clean set of URL is based on a comparison of a fingerprint value calculated for each of the URLs in the clean set of URLs (Office Action, p. 4). This feature is not present in amended claim 1. This section, or any other section, of NAJORK does not disclose or suggest anything about session identifiers. NAJORK does not disclose or suggest analyzing the set of URLs extracted from the at least one document to determine those in the set of URLs that contain session identifiers by locating the session identifiers in the set of URLs extracted as sub-strings that occur in multiple URLs of a web site, as recited in claim 1.

Because GALAI and NAJORK do not disclose or suggest analyzing the set of URLs extracted from the at least one document to determine those in the set of URLs that contain session identifiers by locating the session identifiers in the set of URLs extracted as sub-strings that occur in multiple URLs of a web site, the combination of GALAI and NAJORK does not disclose or suggest analyzing the set of URLs extracted from the at least one document to determine those in the set of URLs that contain session identifiers by locating the session identifiers in the set of URLs extracted as sub-strings that occur in multiple URLs of a web site.

For at least the foregoing reasons, Applicant submits that claim 1 is patentable over GALAI and NAJORK, whether taken alone or in any reasonable combination.

Claims 2-5 and 7-9 depend from claim 1. Therefore, these claims are patentable over GALAI and NAJORK, whether taken alone or in any reasonable combination, for at least the reasons set forth above with respect to claim 1.

Independent claims 10, 15, 20, and 25 recite features similar to, yet possibly of different scope than, the features recited above with respect to claim 1. Therefore, these

claims are patentable over GALAI and NAJORK, whether taken alone or in any reasonable combination, for at least reasons similar to the reasons set forth above with respect to claim 1.

Claims 11-14 depend from claim 10. Therefore, these claims are patentable over GALAI and NAJORK, whether taken alone or in any reasonable combination, for at least the reasons set for the above with respect to claim 10.

Claims 16-19 depend from claim 15. Therefore, these claims are patentable over GALAI and NAJORK, whether taken alone or in any reasonable combination, for at least the reasons set for the above with respect to claim 15.

Claims 21-24 depend from claim 20. Therefore, these claims are patentable over GALAI and NAJORK, whether taken alone or in any reasonable combination, for at least the reasons set for the above with respect to claim 20.

Claims 26-29 depend from claim 25. Therefore, these claims are patentable over GALAI and NAJORK, whether taken alone or in any reasonable combination, for at least the reasons set for the above with respect to claim 25.

In view of the foregoing amendments and remarks, Applicant respectfully requests the Examiner's reconsideration of this application, and the timely allowance of the pending claims.

While the present application is believed to be in condition for allowance, should the Examiner find some issue to remain unresolved, or should any new issues arise that could be eliminated through discussions with Applicant's representative, then the Examiner is invited to contact the undersigned by telephone to expedite prosecution of the present application.

To the extent necessary, a petition for an extension of time under 37 CFR § 1.136 is hereby made. Please charge any shortage in fees due in connection with the filing of this paper, including extension of time fees, to Deposit Account No. 50-1070 and please credit any excess fees to such deposit account.

Respectfully submitted,

**HARRITY SNYDER, L.L.P.**

By: /Viktor Simkovic, Reg. No. 56012/  
Viktor Simkovic  
Registration No. 56,012

Date: January 22, 2008

Harrity Snyder, L.L.P.  
11350 Random Hills Road  
Suite 600  
Fairfax, Virginia 22030  
Main: (571) 432-0800  
Direct: (571) 432-0899  
Customer Number: **44989**